

# ON THE SPREAD OF SUPERCRITICAL RANDOM GRAPHS

LOUIGI ADDARIO-BERRY, SVANTE JANSON, AND COLIN MCDIARMID

**ABSTRACT.** The *spread* of a connected graph  $G$  was introduced by Alon Boppana and Spencer [1] and measures how tightly connected the graph is. It is defined as the maximum over all Lipschitz functions  $f$  on  $V(G)$  of the variance of  $f(X)$  when  $X$  is uniformly distributed on  $V(G)$ . We investigate the spread of a variety of random graphs, in particular the random regular graphs  $G(n, d)$ ,  $d \geq 3$ , and Erdős-Rényi random graphs  $G_{n,p}$  in the supercritical range  $p > 1/n$ . We show that if  $p = c/n$  with  $c > 1$  fixed then with high probability the spread is bounded, and prove similar statements for  $G(n, d)$ ,  $d \geq 3$ . We also prove lower bounds on the spread in the barely supercritical case  $p - 1/n = o(1)$ . Finally, we show that for  $d$  large the spread of  $G(n, d)$  becomes arbitrarily close to that of the complete graph  $K_n$ .

## 1. INTRODUCTION

If  $G$  is a graph, a *Lipschitz function*  $f$  on  $G$  is a real-valued function defined on the vertex set  $V(G)$  such that  $|f(v) - f(w)| \leq 1$  for every pair of adjacent vertices  $v, w$ . We may regard a function  $f : V(G) \rightarrow \mathbb{R}$  on a graph  $G$  as a random variable by evaluating  $f$  at a random, uniformly distributed, vertex. We may thus talk about, e.g., the mean, median and variance of  $f$ . For example, if  $G$  has  $n$  vertices, the mean  $\mathbb{E} f$  is  $\sum_v f(v)/n$ .

For a fixed connected graph  $G$ , we define the *spread*  $\sigma^{2*}(G)$  to be the supremum of  $\text{Var}(f)$  over all Lipschitz functions  $f : V(G) \rightarrow \mathbb{R}$ . (Note that the supremum would be infinite if we considered a disconnected graph.) The spread of a graph was introduced in [1], and considered further in [3]. The purpose of this paper is to investigate the spread for a few types of random graphs.

For every connected graph, the spread is attained, so we can replace supremum by maximum. In fact, it is shown in Theorem 2.1 of [1] that there is always an optimal function  $f$  which is integer-valued and of a simple form. Note that the spread is an edge-monotone function of  $G$  in the sense that if we add an edge to  $G$ , then the set of Lipschitz functions becomes smaller, and thus the spread becomes smaller or remains the same. The

---

*Date:* February 6, 2009.

*2000 Mathematics Subject Classification.* 60C05.

*Key words and phrases.* Random graphs, random regular graphs, spread of a graph, Lipschitz functions.

spread is easily seen to be at most  $(\text{diameter})^2/4$ : our results will imply that it typically is much smaller.

In Section 2 we study random regular graphs with fixed degree  $d \geq 3$ , and show that the spread w.h.p. is bounded by some constant, independent of the size  $n$  of the graph. (We use w.h.p. (*with high probability*) for events with probability  $1 - o(1)$  as  $n \rightarrow \infty$ .)

In Section 3 we study the random graph  $G_{n,c/n}$  with fixed  $c > 1$ . This random graph is w.h.p. disconnected: so we consider the spread of the largest component of  $G_{n,c/n}$ . We denote the largest component of  $G_{n,p}$  by  $H_{n,p}$ , and show that the spread  $\sigma^{2*}(H_{n,c/n})$  is w.h.p. bounded by a constant depending on  $c$ . (Recall that for  $c > 1$ , there is w.h.p. a unique giant component  $H_{n,c/n}$  of order  $\sim \gamma(c)n$  for some  $\gamma(c) > 0$ .)

In Section 4 we study the random graph  $G_{n,c/n}$  in the barely supercritical case when  $c = 1 + \epsilon$ ,  $\epsilon = \epsilon(n) \rightarrow 0$  but  $\epsilon^3 n \rightarrow \infty$  as  $n \rightarrow \infty$ ; we show that in this case the spread tends to infinity (in probability), at least at the rate  $\epsilon^{-2}$ .

In Section 5 we study the random regular graph  $G(n, d)$  in the large- $d$  case. It is easily seen by comparison with the complete graph  $K_n$  that the spread of  $G(n, d)$  always is at least  $1/4 - 1/(4n^2)$ . We show that for any  $\epsilon > 0$ , for all  $d$  sufficiently large w.h.p. the spread of  $G(n, d)$  is at most  $1/4 + \epsilon$ .

Section 6 contains some open problems, inspired by our results.

We use  $c_1, C_1$  etc. to denote various positive constants. (We use  $c_i$  for small constants and  $C_i$  for large.) In Section 3, where we consider  $G_{n,c/n}$ , these are allowed to depend on  $c$ , but they never depend on  $n$ .

**Acknowledgement.** This work was started during the Oxford–Princeton workshop on Combinatorics in Oxford, June 2008, and continued during visits of CM and SJ to CRM, Montreal in September and October 2008.

## 2. RANDOM REGULAR GRAPHS

$G(n, d)$  is the random regular graph with degree  $d$ . (If  $d$  is odd,  $n$  is required to be even.)

**Theorem 2.1.** *Fix  $d \geq 3$ . There exists a constant  $c_1$  such that w.h.p.  $G(n, d)$  is such that every Lipschitz function  $f : G(n, d) \rightarrow \mathbb{R}$  satisfies*

$$|\{v : |f(v) - m| \geq x\}| < e^{-c_1 x} n, \quad x \geq 1, \quad (2.1)$$

where  $m$  is a median of  $f$ .

**Corollary 2.2.** *There exists a constant  $C_1$  such that, for every fixed  $d \geq 3$ , w.h.p.  $\sigma^{2*}(G(n, d)) \leq C_1$ , i.e.,  $G(n, d)$  is w.h.p. such that every Lipschitz function  $f : G(n, d) \rightarrow \mathbb{R}$  satisfies  $\text{Var}(f) \leq C_1$ .*

In principle, numerical values could be given for the constants  $c_1$  and  $C_1$ , but we have not tried to find explicit values, nor to optimize the arguments. These constants can be taken independent of  $d \geq 3$ ; in fact, it follows by monotonicity [8, Theorem 9.36] that any constant that works for  $d = 3$  will

work for all larger  $d$  as well. We will thus consider  $d = 3$  only in the proof. (Alternatively, and perhaps more elementarily, we are convinced that the proof below easily could be modified to an arbitrary  $d$ , but we have not checked the details.)

For  $\alpha > 0$  we say that a graph  $G$  is an  $\alpha$ -*expander* if every set  $W \subset V(G)$  with  $|W| \leq |V(G)|/2$  contains at least  $\alpha|W|$  vertices with neighbours in  $V(G) \setminus W$ . (This is slightly at odds with the standard definition of expansion but is more convenient for our purposes.) The *Cheeger constant* of  $G$  is

$$\Phi(G) = \min_{\{S \subset V(G) : \sum_{v \in V(S)} d(v) \leq |E(G)|\}} \frac{|E(S, V(G) \setminus S)|}{\sum_{v \in V(S)} d(v)}. \quad (2.2)$$

$\Phi(\cdot)$  measures the edge expansion, rather than the vertex expansion, of graphs. We shall use the following expander property of  $G(n, 3)$ , proved (in a more general version) by [2] (see also [10] and [6, (proof of) Lemma 5.1]).

**Lemma 2.3** ([2], Lemma 5.3). *There exists  $c_2 > 0$  such that w.h.p.  $\Phi(G(n, 3)) \geq c_2$ .*

Since  $G(n, 3)$  has constant degree, Lemma 2.3 immediately implies vertex expansion for  $G(n, 3)$ , with the same constant. We state this as a simple lemma.

**Lemma 2.4.** *If  $G$  is regular, and  $0 < \alpha \leq \Phi(G)$ , then  $G$  is an  $\alpha$ -expander.*

*Proof.* Let  $n := |G|$ , and let  $d$  be the degree of the vertices. Note that  $G$  has precisely  $dn/2$  edges. Fix a set  $W$  of vertices in  $G$  with  $|W| \leq n/2$ . Then  $\sum_{v \in W} d(v) = d|W| \leq dn/2$ , so by (2.2) there are at least  $\Phi(G)d|W|$  edges from  $W$  to its complement. These edges have at least  $\Phi(G)d|W|/d \geq \alpha|W|$  endpoints in  $W$ .  $\square$

**Lemma 2.5.**  *$G(n, 3)$  is w.h.p. a  $c_2$ -expander.*

*Proof.* An immediate consequence of Lemmas 2.3 and 2.4.  $\square$

*Proof of Theorem 2.1.* As said above, it suffices to prove the result for  $d = 3$ . Assume that the conclusion of Lemma 2.5 holds, and that  $G(n, 3)$  is connected; both these hold w.h.p. by Lemma 2.5 and [5; 16], see also [4].

Let  $f$  be a Lipschitz function on  $G(n, 3)$ , and let  $m$  be its median. We may assume that  $m = 0$ ; otherwise we replace  $f$  by  $f - m$ . Let  $E_t := \{v \in [n] : f(v) \geq t\}$ . Then  $|E_t| \leq n/2$  for  $t > 0$ .

If  $t > 0$  and  $E_t$  is nonempty then by our assumptions on the graph  $G(n, 3)$ , there is a subset of  $E_t$  of size at least  $c_2|E_t|$  of vertices  $x$  that are adjacent to at least one vertex  $y \notin E_t$ . Thus  $f(y) < t$ , and since  $f$  is Lipschitz, we have  $f(x) < t + 1$  for every such  $x$ . Consequently,  $|E_{t+1}| \leq (1 - c_2)|E_t|$  when  $t > 0$ . Since  $|E_1| \leq n/2 \leq (1 - c_2)n$  (assuming as we may that  $c_2 \leq 1/2$ ), we obtain by induction, for simplicity considering integers  $k$  only,

$$|E_k| \leq (1 - c_2)^k n \leq e^{-c_2 k} n, \quad k = 1, 2, \dots \quad (2.3)$$

By symmetry, we have the same estimate for  $\{v : f(v) \leq -k\}$ , and thus, for every  $x \geq 1$ ,

$$|\{v : |f(v)| \geq x\}| \leq 2e^{-c_2 \lfloor x \rfloor} n < 2e^{-(c_2/2)x} n. \quad (2.4)$$

For  $x \geq C_2$ , this estimate is less than  $e^{-(c_2/3)x} n$ , and it is easy to see that the result holds trivially for smaller  $x$  too if we choose  $c_1 < c_2/3$  small enough.  $\square$

### 3. $G_{n,c/n}$ WITH $c > 1$ FIXED.

For  $G_{n,c/n}$  with  $c > 1$  fixed, in place of Lemmas 2.3 and 2.5 we can use another result of [2]. For a graph  $G$  and a set of vertices  $U \subset V(G)$ , we write  $G \setminus U$  for the subgraph of  $G$  induced by  $V(G) \setminus U$ . For  $0 < \alpha < 1$  we say that a connected graph  $G$  is an  $\alpha$ -decorated expander if  $G$  has a subgraph  $F$  such that

(DE1)  $\Phi(F) \geq \alpha$ ;

(DE2) listing the connected components of  $G \setminus V(F)$  as  $D_1, \dots, D_\ell$  for some  $\ell$ , for  $x > 0$  we have

$$|\{i : |E(D_i)| + |E(D_i, F)| \geq x\}| \leq e^{-\alpha x} \cdot |E(G)|;$$

(DE3) no vertex  $v \in V(F)$  is adjacent to (“decorated by”) more than  $1/\alpha$  of the components  $D_i$ .

Note that (DE1) implies that  $F$  is connected. Note further that (DE2) implies:

(DE2') for all  $x \geq 0$ ,  $|\{i : |V(D_i)| \geq x\}| \leq e^{-\alpha x} \cdot |E(G)|$ .

We shall use (DE2') rather than (DE2) in what follows. Benjamini et. al. [2] prove the following (their Theorem 4.2 and Lemma 4.7, combined):

**Theorem 3.1.** *Fix  $c > 1$ . Then there is a constant  $\alpha = \alpha(c) > 0$  such that w.h.p. the giant component  $H_{n,c/n}$  of  $G_{n,c/n}$  is an  $\alpha$ -decorated expander. More strongly, w.h.p.  $H_{n,c/n}$  has a connected subgraph  $F$  with at least  $\alpha n$  vertices satisfying conditions (DE1)–(DE3) in the definition of  $\alpha$ -decorated expanders.*

(To be precise, (DE2) is stated in [2] with  $|E(G_{n,c/n})|$  rather than  $|E(H_{n,c/n})|$ ; this is equivalent, by a change of  $\alpha$ , since w.h.p. the giant component contains at least a fixed fraction of all edges.)

Since the expansion guaranteed by Theorem 3.1 is edge-expansion, we will need to do a little work to derive the vertex expansion required to prove Theorem 3.3, below. The following lemma will give some further, more elementary, properties of  $G_{n,c/n}$  that suffice for our purposes. Let  $V_i = V_i(n)$  be the set of vertices of degree  $i$  in  $G_{n,c/n}$ .

The constants  $C_3, C_4, \dots$  below may depend on  $c$ .

**Lemma 3.2.**  *$G_{n,c/n}$  is w.h.p. such that the following properties hold:*

(P1)  $|E(G_{n,c/n})| \leq cn$ ,

(P2)  $n' := |V(H_{n,c/n})| > \gamma n/2$  for some  $\gamma = \gamma(c) > 0$ ,

(P3)  $|V_i| \leq ne^{-i}$  for all  $i \geq C_3$ .

*Proof.* It is well-known and easy to see that  $|E(G_{n,c/n})|/n \xrightarrow{P} c/2$ , which entails (P1). It is also well-known that  $n'/n \xrightarrow{P} \gamma(c) > 0$ , which entails (P2).

For (P3), let  $d_j$  be the degree of vertex  $j$ , and let  $X$  be the random variable  $\sum_{j=1}^n e^{2d_j}$ . Since each  $d_j$  has a binomial  $\text{Bin}(n-1, c/n)$  distribution,

$$\mathbb{E} X = n \mathbb{E} e^{2d_1} = n \left(1 + \frac{c}{n}(e^2 - 1)\right)^{n-1} \leq ne^{c(e^2-1)}. \quad (3.1)$$

A similar calculation shows that  $\text{Var } X = n \text{Var}(e^{d_1}) + n(n-1) \text{Cov}(e^{d_1}, e^{d_2}) = O(n)$ . Consequently, by Chebyshev's inequality, w.h.p.

$$\sum_{i=0}^{\infty} e^{2i} |V_i| = X \leq e^{ce^2} n.$$

The result follows.  $\square$

**Theorem 3.3.** *Given fixed  $c > 1$  there exist constants  $C_4 > 0$  and  $C_5 > 1$  such that w.h.p. the giant component  $H_{n,c/n}$  of  $G_{n,c/n}$  is such that every Lipschitz function  $f : H_{n,c/n} \rightarrow \mathbb{R}$  satisfies*

$$|\{v : |f(v) - m| > x\}| < C_4 C_5^{-\sqrt{x}} n, \quad x \geq 0, \quad (3.2)$$

where  $m$  is a median of  $f$ .

**Corollary 3.4.** *For fixed  $c > 1$  there exists a constant  $C_6 = C_6(c) > 0$  such that w.h.p.  $\sigma^{2*}(G_{n,c/n}) \leq C_6$ , i.e., w.h.p.  $G_{n,c/n}$  is such that every Lipschitz function  $f : H_{n,c/n} \rightarrow \mathbb{R}$  satisfies  $\text{Var}(f) \leq C_6$ .*

Theorem 3.3 follows immediately from Theorem 3.1, Lemma 3.2 and the following deterministic lemma. Corollary 3.4 is an immediate consequence.

**Lemma 3.5.** *Suppose that  $G_{n,c/n}$  and  $H_{n,c/n}$  satisfy the properties in Theorem 3.1 and Lemma 3.2. Then (3.2) holds for every Lipschitz function  $f$  on  $H_{n,c/n}$ .*

*Proof.* Fix a subgraph  $F$  of  $H = H_{n,c/n}$  as described in Theorem 3.1; let  $D_1, \dots, D_\ell$  be the components of  $H \setminus V(F)$ ; let  $D$  be the graph with components the  $D_i$ ; and fix a Lipschitz function  $f$  on the vertices of  $H$ . Let  $n' = |V(H)|$  as in Lemma 3.2.

We write  $H_{\geq t}$  for the set of vertices  $v \in V(H)$  with  $f(v) \geq t$  and define  $H_{>t}$ ,  $H_{\leq t}$ ,  $H_{<t}$  similarly. Additionally, we write  $F_{\geq t}$  (and  $F_{>t}$  et cetera) for  $V(F) \cap H_{\geq t}$ , and  $D_{\geq t}$  (et cetera) for  $H_{\geq t} \cap V(D) = H_{\geq t} \setminus V(F)$ . We also assume as in the proof of Theorem 2.1 that  $f$  has median  $m = 0$ ; hence  $|H_{\leq 0}|, |H_{\geq 0}| \geq n'/2$ .

Our plan of attack is as follows. First, we find a large subset of  $V(F)$  consisting exclusively of vertices  $v$  with  $f(v) = O(1)$ . Such a set is not quite guaranteed by the fact that  $|H_{\leq 0}| \geq n'/2$ , because  $H_{\leq 0}$  may be largely contained within  $V(H) \setminus V(F)$ . However, we can use properties (DE2') and

(DE3) to find such a set. Second, we use the expansion of  $F$  to show that the sets  $F_{\geq t}$  decay rapidly in size as  $t$  grows. Finally, we use the fact that the decorations  $D_i$  are typically small and do not attach to very many vertices of  $F_{\geq t}$ , to show that the sets  $D_{\geq t}$  also decay rapidly in size as  $t$  grows. We now turn to the details. For simplicity we prove the theorem for  $x$  integer, which easily implies the more general statement.

For  $\lambda > 0$  let  $F^\lambda$  be the union of  $F$  and all components  $D_i$  containing less than  $\lambda$  vertices. By property (P1),  $|E(H)| \leq |E(G_{n,c/n})| \leq cn$ ; this is a very crude but sufficient bound that we will use whenever we need to control  $|E(H)|$ , and accounts for essentially all the subsequent occurrences of the constant  $c$ .

By property (DE2'), for any  $\lambda > 0$  we have

$$\begin{aligned} \sum_{\{i : |V(D_i)| \geq \lambda\}} |V(D_i)| &\leq \sum_{j=0}^{\infty} \sum_{\{i : 2^j \lambda \leq |V(D_i)| < 2^{j+1} \lambda\}} |V(D_i)| \\ &\leq \sum_{j=0}^{\infty} 2^{j+1} \lambda e^{-\alpha \lambda 2^j} \cdot cn. \end{aligned} \quad (3.3)$$

Choose  $\lambda = \lambda_1$  large enough that the upper bound in (3.3) is less than  $\gamma n/8 < n'/4$ , using also property (P2); then  $F^{\lambda_1}$  contains at least  $3n'/4$  vertices. Since at most  $n'/2$  vertices  $v$  in  $H$  have  $f(v) > 0$ , it follows that at least  $n'/4$  of the vertices in  $F^{\lambda_1}$  have  $f(v) \leq 0$ . Since each component of  $F^{\lambda_1} \setminus F$  has less than  $\lambda_1$  vertices, either  $|F_{\leq 0}| \geq n'/8$  or at least  $n'/(8\lambda_1)$  components of  $F^{\lambda_1} \setminus F$  contain a vertex of  $H_{\leq 0}$ . Since all vertices in  $F^{\lambda_1} \setminus F$  have distance at most  $\lambda_1$  from  $F$ , property (DE3) and the Lipschitz property of  $f$  then guarantee that in either case (assuming  $\lambda_1 > \alpha$  as we may)

$$|F_{\leq \lambda_1}| \geq \frac{\alpha n'}{8\lambda_1} =: c_3 n'.$$

Since every vertex of  $F$  has at least one neighbour in  $F$ , it follows that  $\sum_{v \in F_{\leq \lambda_1}} d_F(v) \geq c_3 n'$ . Assuming that  $\sum_{v \in F_{\leq \lambda_1}} d_F(v) \leq |E(F)|$ , by the expansion property (DE1) we thus have that  $|E(F_{\leq \lambda_1}, F_{> \lambda_1})| \geq \alpha c_3 n'$ . The Lipschitz property of  $f$  implies that each edge in  $E(F_{\leq \lambda_1}, F_{> \lambda_1})$  has one endpoint in  $F_{\leq \lambda_1+1} \setminus F_{\leq \lambda_1}$ , and thus using again property (P2),

$$\sum_{v \in F_{\leq \lambda_1+1} \setminus F_{\leq \lambda_1}} d_F(v) \geq |E(F_{\leq \lambda_1}, F_{> \lambda_1})| \geq \alpha c_3 n' \geq (\alpha c_3 \gamma/2)n =: c_4 n.$$

Repeatedly applying property (DE1) in this manner, and using property (P2), we see that w.h.p.  $\sum_{v \in F_{\leq \lambda_2}} d_F(v) \geq |E(F)|$ , where we may take  $\lambda_2 = \lambda_1 + c/c_4 + 1$ .

Next, choose an integer  $C_7 \geq C_3$  large enough that  $\sum_{i=C_7+1}^{\infty} i e^{-i} \leq \alpha/4$ ; hence by property (P3), we then have  $\sum_{v \in F} d_F(v) \mathbf{1}[d_F(v) > C_7] \leq \alpha n/4$ .

Since  $|E(F)| \geq |V(F)| - 1 \geq \alpha n - 1 \geq \alpha n/2$  (for  $n > 2/\alpha$ ), it follows that

$$\sum_{v \in F_{\leq \lambda_2}} d_F(v) \mathbf{1}[d_F(v) \leq C_7] \geq |E(F)| - \alpha n/4 \geq |E(F)|/2,$$

and so

$$|F_{\leq \lambda_2}| \geq |E(F)|/(2C_7) \geq \alpha n/(4C_7) =: c_5 n. \quad (3.4)$$

We next apply the expansion of  $F$  and properties (P1)–(P3) to bound the sizes of sets  $F_{>\lambda_2+i}$  for positive integers  $i$ . As  $i$  becomes large and the sets  $F_{>\lambda_2+i}$  become small, the proportion of the sum  $\sum_{v \in F_{>\lambda_2+i}} d_F(v)$  due to vertices of large degree may increase; this is the reason we are only able to show that the sizes of the sets  $F_{>\lambda_2+i}$  decay exponentially quickly in  $\sqrt{i}$ .

For given  $x > 0$ , let  $a_x$  be the smallest integer  $\geq C_3$  such that  $\sum_{i>a_x}^{\infty} i e^{-i} \leq \alpha x/2$ . Since  $\sum_{i>a}^{\infty} i e^{-i} \leq \sum_{i>a}^{\infty} e^{-i/2} \leq 3e^{-a/2}$ , there exists  $C_8$  large enough that  $a_x \leq C_8 \log_2(1/x)$  for all  $x \leq 1/2$ .

For  $\lambda \geq \lambda_2$ , if  $t' = \sum_{v \in F_{>\lambda}} d_F(v)$  then  $t' \leq |E(F)|$  by our choice of  $\lambda_2$ . For any  $t \leq \min(t', n/2)$ , we thus have  $|E(F_{>\lambda}, F_{\leq \lambda})| \geq \alpha t' \geq \alpha t$  by (DE1). Let

$$\partial F_{>\lambda} = \{v \in F_{>\lambda} : v \text{ has a neighbour in } F_{\leq \lambda}\}.$$

Then for any  $t$  as above,  $\sum_{v \in \partial F_{>\lambda}} d_F(v) \geq |E(F_{>\lambda}, F_{\leq \lambda})| \geq \alpha t$ , and applying property (P3) as previously we have  $\sum_{v \in \partial F_{>\lambda}} d_F(v) \mathbf{1}[d_F(v) \leq a_{t/n}] \geq \alpha t/2$ , so

$$|\partial F_{>\lambda}| \geq \frac{\alpha t/2}{a_{t/n}} \geq \frac{\alpha t}{2C_8 \log_2(n/t)} := c_6 \frac{t}{\log_2(n/t)}.$$

Now fix  $\lambda \geq \lambda_2$ . Taking  $t = |F_{>\lambda}| \leq \sum_{v \in F_{>\lambda}} d_F(v) = t'$ , we also have  $t \leq |H_{>0}| \leq n/2$ , so the preceding inequality applies with this choice of  $t$ . Furthermore, the Lipschitz property of  $f$  implies that  $\partial F_{>\lambda} \subseteq F_{\leq \lambda+1}$ , and so

$$|F_{>\lambda+1}| \leq |F_{>\lambda}| - |\partial F_{>\lambda}| \leq t(1 - c_6/\log_2(n/t)).$$

Next, for integers  $i \geq 1$ , let  $k_i = \lceil i/c_6 \rceil$ . Then for all  $t \geq n/2^i$ , we have  $(1 - c_6/\log_2(n/t))^{k_i} < 1/2$ . It follows immediately that for all integers  $i \geq 0$  we have

$$|F_{>\lambda_2 + \sum_{j=1}^i k_j}| \leq \frac{n}{2^{i+1}},$$

so there is  $A > 0$  such that for all real  $x \geq 0$ ,

$$|F_{>Ax^2}| \leq \frac{n}{2^x}. \quad (3.5)$$

We now deal with the elements of the ‘decorations’ graph  $D$ , and assume that its components  $D_1, \dots, D_\ell$  are listed in decreasing order of number of vertices. We first remark that by (DE2') and (P1), if  $m_k$  is the number of components of  $D$  of size at least  $k$ , then  $m_k \leq c n e^{-\alpha k}$  for all integers  $k \geq 1$ .

Hence, for any real  $t$  with  $0 < t \leq n$ , we have, with  $x = \log(cn/t)/\alpha$ ,

$$\begin{aligned} \sum_{j=1}^{\lfloor t \rfloor} |V(D_j)| &= \sum_{k=1}^{\infty} \min(\lfloor t \rfloor, m_k) \leq \sum_{k=1}^{\infty} \min(t, cne^{-\alpha k}) \\ &= \sum_{k \leq x} t + \sum_{k > x} cne^{-\alpha k} \leq C_9 t (\log n + 1 - \log t). \end{aligned} \quad (3.6)$$

Next, for  $v \in D$ , let  $D(v)$  be the component of  $D$  containing  $v$  and let  $u = u(v)$  be the vertex of  $F$  decorated with  $D(v)$  and with  $f(u)$  maximal subject to this (with ties broken arbitrarily). By (DE3), for any set  $S \subseteq V(F)$  with  $|S| \leq s$ , the total number of components that decorate  $S$  is at most  $s/\alpha$ . It then follows from (3.6) that

$$|\{v \in D : u(v) \in S\}| \leq \sum_{j=1}^{\lfloor s/\alpha \rfloor} |V(D_j)| \leq C_{10} s (\log n + 1 - \log s) \quad (3.7)$$

if  $s \leq \alpha n$ , and by taking  $C_{10} \geq 1/\alpha$  we see that the inequality in fact holds for all  $s \leq n$ .

For  $i \geq 0$ , if  $v \in D_{>i}$  then one of the following two events must occur.

- (a)  $|D(v)| \geq 3i/4$ .
- (b)  $u(v) \in F_{>i/4}$ .

By (DE2'),

$$|\{v \in D : |D(v)| \geq 3i/4\}| \leq \sum_{j \geq 3i/4} j \cdot cne^{-\alpha j} \leq C_{11} nie^{-3\alpha i/4}. \quad (3.8)$$

Furthermore, by (3.5),

$$|F_{>i/4}| \leq n/2^{c_7 \sqrt{i}}$$

and thus by (3.7) we have

$$|\{v \in D : u(v) \in F_{>i/4}\}| \leq C_{10} \frac{n}{2^{c_7 \sqrt{i}}} (1 + c_7 \sqrt{i} \log 2),$$

so for all  $i$  we have

$$|\{v \in D : u(v) \in F_{>i/4}\}| \leq C_{12} \frac{n}{C_5^{\sqrt{i}}} \quad (3.9)$$

for suitable constants  $C_{12}$  and  $C_5 > 1$ . Thus, by (3.8) and (3.9),

$$|D_{>i}| \leq |\{v \in D : |D(v)| \geq 3i/4\}| + |\{v \in D : u(v) \in F_{>i/4}\}| \leq C_{13} \frac{n}{C_5^{\sqrt{i}}}. \quad (3.10)$$

Hence, using this together with (3.5) to bound  $|F_{>i}|$ , we have

$$|H_{>i}| = |F_{>i}| + |D_{>i}| \leq C_{14} C_5^{-\sqrt{i}} n$$

for fixed  $C_{14}$  sufficiently large. To complete the proof, note that  $-f$  is also a Lipschitz function on  $H$  with a median 0, and so we may set  $C_4 = 2C_{14}$ .  $\square$



4.  $G_{n,(1+\epsilon)/n}$  WITH  $\epsilon \rightarrow 0$ ,  $\epsilon \gg n^{-1/3}$ .

Fix a function  $\epsilon = \epsilon(n)$  as above and let  $p = (1 + \epsilon)/n$ . As above, denote by  $H_{n,p}$  the largest component of  $G_{n,p}$ . Additionally, write  $C_{n,p}$  (resp.  $K_{n,p}$ ) for the core (resp. kernel) of  $H_{n,p}$ . For such  $\epsilon$ , it is known (see [12] and also [8], Chapter 5) that w.h.p.

$$\begin{aligned} |V(H_{n,p})| &= (1 + o(1))2\epsilon n, \\ |V(C_{n,p})| &= (1 + o(1))2\epsilon^2 n, \text{ and} \\ |V(K_{n,p})| &= (1 + o(1))\frac{4}{3}\epsilon^3 n. \end{aligned} \tag{4.1}$$

For a connected graph  $G$ , we write  $\kappa(G) = |E(G)| - |V(G)|$ , and call  $\kappa$  the *excess* of  $G$ . A moment's reflection reveals that  $\kappa(H_{n,p}) = \kappa(C_{n,p}) = \kappa(K_{n,p})$ , and it is known ([7; 8; 11]) that for  $\epsilon$  as above, w.h.p.

$$\kappa(H_{n,p}) = (1 + o(1))\frac{2}{3}\epsilon^3 n. \tag{4.2}$$

We fix  $\delta < 1/10$  and say  $H_{n,p}$  *behaves* if

$$(2 - \delta)\epsilon n \leq |V(H_{n,p})| \leq (2 + \delta)\epsilon n,$$

and if similar inequalities hold for  $|V(C_{n,p})|$ ,  $|V(K_{n,p})|$ , and  $\kappa(H_{n,p})$ . By the above comments, w.h.p.  $H_{n,p}$  behaves. Using this fact, we prove the following theorem.

**Theorem 4.1.** *Let  $p = (1 + \varepsilon)/n$  with  $\varepsilon = \varepsilon(n) = o(1)$  and  $\varepsilon^3 n \rightarrow \infty$  as  $n \rightarrow \infty$ . Then w.h.p.  $\sigma^{2*}(H_{n,p}) = \Omega(1/\varepsilon^2)$ .*

The complement  $H_{n,p} \setminus V(C_{n,p})$  of the core in the largest component  $H_{n,p}$  is a forest consisting of trees that are attached to the core by (exactly) one edge each. We call these trees *pendant*, and denote them (in some order) by  $T_1, \dots, T_N$ . We begin with an estimate of the maximum size of the pendant trees.

**Lemma 4.2.** *There exists a constant  $C_{15}$  such that w.h.p.*

$$\max_i |V(T_i)| \leq C_{15}\varepsilon^{-2} \log(n\varepsilon^3). \tag{4.3}$$

*Proof.* We create another forest by removing all edges in the core  $C_{n,p}$  from  $H_{n,p}$ ; the result is a forest where each component consists of a single vertex in  $V(C_{n,p})$  together with all pendant trees attached to it (if any). We regard these trees as rooted, with the vertices in  $V(C_{n,p})$  as the roots, and denote them by  $T_v^*$ ,  $v \in V(C_{n,p})$ .

Conditioned on  $V(H_{n,p})$  and  $C_{n,p}$ , this forest  $\{T_v^*\}_v$  is a uniformly distributed forest of rooted trees, with given sets of  $M := |V(C_{n,p})|$  roots and  $m := |V(H_{n,p})| - M$  non-roots.

The maximum size of a tree in a random forest of rooted trees has been studied by Pavlov [13] (see also [9, Section 3.6] and [14]). In our case we have, if  $H_{n,p}$  behaves and  $n$  is large enough,  $(2 - \delta)n\varepsilon^2 \leq M \leq (2 + \delta)n\varepsilon^2$

and  $(2 - 2\delta)n\varepsilon \leq m \leq (2 + \delta)n\varepsilon$ . In particular,  $m/M \rightarrow \infty$  and  $m/M^2 \leq (n\varepsilon^3)^{-1} \rightarrow 0$ . This is the range of [13, Theorem 3 (and the remark following it)], which implies that w.h.p., conditioned on  $M$  and  $m$ ,

$$\max_v |V(T_v^*)| = (1 + o(1)) \frac{2m^2}{M^2} \log \left( \frac{M^2}{m} \right) \leq C_{15}\varepsilon^{-2} \log(n\varepsilon^3).$$

The same estimate thus holds unconditionally w.h.p., and the result follows since every pendant tree is contained in some  $T_v^*$ .  $\square$

*Proof of Theorem 4.1.* Since  $H_{n,p}$  behaves w.h.p., it suffices to prove that *given* that  $H_{n,p}$  behaves, w.h.p.  $\sigma^{2*}(H_{n,p}) = \Omega(1/\varepsilon^2)$ . We shall define a Lipschitz function  $f$  on the vertices of  $H_{n,p}$  for which, given that  $H_{n,p}$  behaves, w.h.p.  $\text{Var}(f) \geq \gamma/\varepsilon^2$  for some fixed  $\gamma > 0$ . We define  $f$  in a few steps, starting from the core. We say that  $e \in E(K_{n,p})$  has *length*  $\ell(e)$  if the path in  $C_{n,p}$  corresponding to  $e$  contains  $\ell(e)$  edges (so  $\ell(e) - 1$  internal vertices). Since  $H_{n,p}$  behaves,

$$\begin{aligned} |E(K_{n,p})| &= |V(K_{n,p})| + \kappa(K_{n,p}) \\ &\leq \left(\frac{4}{3} + \delta\right)\varepsilon^3 n + \left(\frac{2}{3} + \delta\right)\varepsilon^3 n \\ &= (2 + 2\delta)\varepsilon^3 n, \end{aligned} \tag{4.4}$$

and

$$\begin{aligned} |V(C_{n,p}) \setminus V(K_{n,p})| &\geq (2 - \delta)\varepsilon^2 n - \left(\frac{4}{3} + \delta\right)\varepsilon^3 n \\ &\geq (2 - 2\delta)\varepsilon^2 n, \end{aligned} \tag{4.5}$$

for  $n$  sufficiently large.

We say that an edge  $e \in E(K_{n,p})$  is *short* if

$$\ell(e) \leq \left\lfloor \frac{1 - \delta}{2\varepsilon(1 + \delta)} \right\rfloor$$

(and is long otherwise), and that  $v \in V(C_{n,p}) \setminus V(K_{n,p})$  is *useless* if it is contained in a path corresponding to a short edge (and is useful otherwise). By (4.4) and (4.5), the number of useful vertices is at least

$$|V(C_{n,p}) \setminus V(K_{n,p})| - |E(K_{n,p})| \cdot \frac{1 - \delta}{2\varepsilon(1 + \delta)} \geq (1 - \delta)\varepsilon^2 n. \tag{4.6}$$

Next, let  $r = r(n)$  be the largest integer divisible by 3 and with  $2r \leq (1 - \delta)/(2\varepsilon(1 + \delta))$ . For each long edge  $e \in E(K_{n,p})$ , let  $P_e$  be the path in  $C_{n,p}$  corresponding to  $e$  (so the endpoints of  $P_e$  are in  $K_{n,p}$ ), and let  $P'_e$  be a sub-path of  $P_e$ , not containing the endpoints of  $P_e$ , which is as long as possible subject to the condition that  $2r$  divides  $|V(P'_e)|$  (picked according to some rule); such a sub-path certainly exists since

$$|V(P_e)| = |E(P_e)| + 1 \geq \left\lfloor \frac{1 - \delta}{2\varepsilon(1 + \delta)} \right\rfloor + 2 \geq 2r + 2,$$

so  $P_e$  has at least  $2r$  internal vertices. Since  $P_e$  has  $|V(P_e)| - 2$  internal vertices, we also have that  $|V(P'_e)| \geq (|V(P_e)| - 2)/2$ , so by (4.4) and (4.6),

$$\begin{aligned} |\{v : v \in P'_e \text{ for some } e \in E(K_{n,p})\}| &\geq \sum_{\{e: e \text{ is long}\}} \frac{|V(P_e)| - 2}{2} \\ &\geq \frac{(1 - \delta)\epsilon^2 n}{2} - 2(1 + \delta)\epsilon^3 n \\ &\geq \frac{(1 - \delta)\epsilon^2 n}{3}, \end{aligned} \quad (4.7)$$

for  $n$  large enough. We now define the restriction of  $f$  to  $V(C_{n,p})$  as follows.

- If  $v \in V(K_{n,p})$ ,  $v$  is useless, or  $v$  is not in  $P'_e$  for any long edge  $e$ , then set  $f(v) = 0$ .
- For each long edge  $e$ , repeat the sequence of values  $12 \dots (r-1)rr(r-1) \dots 1$  along  $P'_e$  (so if  $v$  is the  $i$ 'th or  $(2r+1-i)$ 'th vertex  $\bmod 2r$  along some path  $P'_e$  then  $f(v) = i$ ).

To extend  $f$  from  $C_{n,p}$  to the remainder of  $H_{n,p}$ , for each vertex  $v \in V(H_{n,p})$ , we define the *point of attachment*  $a(v)$  to be the vertex  $x \in C_{n,p}$  whose distance from  $v$  in  $H_{n,p}$  is minimum, and we set  $f(v) = f(a(v))$ . In other words, for each pendant tree  $T$  in  $H_{n,p}$  that hooks up to the core at  $v \in V(C_{n,p})$ , we set  $f(w) = f(v)$  for all  $w \in V(T)$ .

To analyze the variance of  $f$ , for  $i = 1, 2, 3$ , let

$$B_i = \left\{ v \in V(C_{n,p}) : \frac{i-1}{3}r < f(v) \leq \frac{i}{3}r \right\},$$

and let  $B_0$  be all remaining vertices of  $C_{n,p}$ , i.e., those with  $f(v) = 0$ . By the definition of  $f$  and since 3 divides  $r$ , the sizes of  $B_1, B_2$ , and  $B_3$  are identical, and are each at least  $(1 - \delta)\epsilon^2 n/9$ . Also, for  $i = 1, 2, 3$ , let  $B_i^+$  be the set of vertices  $v \in V(H_{n,p})$  with  $a(v) \in B_i$ . We will prove the following assertion:

( $\star$ ) given that  $H_{n,p}$  behaves, w.h.p.  $|B_i^+| \geq \epsilon n/44$  for each  $i = 1, 2, 3$ .

Assuming for the moment that ( $\star$ ) holds, we can quickly complete the proof of the theorem. For each graph  $H_{n,p}$  which behaves, the corresponding (fixed) function  $f$  satisfies

$$\begin{aligned} \text{Var}(f) &= \frac{1}{2(n')^2} \sum_{x, y \in V(H_{n,p})} (f(x) - f(y))^2 \\ &\geq (n')^{-2} \sum_{x \in B_1^+} \sum_{y \in B_3^+} (f(x) - f(y))^2 \\ &\geq (n')^{-2} |B_1^+| |B_3^+| r^2/9 \\ &\geq \frac{(\epsilon n/44)^2}{((2 + \delta)\epsilon n)^2} \frac{r^2}{9} \end{aligned}$$

$$= \frac{r^2}{69696(1 + \delta/2)^2}.$$

But  $r = \Omega(1/\epsilon)$ , and so it follows that, conditional on the event that  $H_{n,p}$  behaves, w.h.p.  $\text{Var}(f) = \Omega(\epsilon^{-2})$ , as needed.

It thus remains to prove  $(\star)$ , and we now turn to this. Let  $X = |B_1^+|$ , the number of vertices  $v \in V(H_{n,p})$  with  $a(v) \in B_1$ . Our aim is to show that  $\mathbb{P}\{X \geq \epsilon n/44\} = 1 - o(1)$ .

We note that given  $C_{n,p}$ , we can specify  $H_{n,p}$  by listing the pendant subtrees of  $H_{n,p}$ , and their points of attachment in  $C_{n,p}$ , as  $T_1, \dots, T_N$  and  $v_1, \dots, v_N$ . By routine calculation it is easily seen that given  $C_{n,p}$  and the pendant subtrees  $T_1, \dots, T_N$ , the points of attachment  $v_1, \dots, v_N$  are independent and uniformly random elements of  $V(C_{n,p})$ . We further note that given  $C_{n,p}$  and the pendant subtrees  $T_1, \dots, T_N$ , we can determine whether or not  $H_{n,p}$  behaves. Then, recalling Lemma 4.2,

$$\mathbb{P}\{X \geq \epsilon n/44\} \geq \inf_{\mathcal{S}} \mathbb{P}\{X \geq \epsilon n/44 \mid C_{n,p}, T_1, \dots, T_N\} - o(1), \quad (4.8)$$

where  $\mathcal{S}$  represents all possible choices of  $C_{n,p}$  and  $N$  and  $T_1, \dots, T_N$  for which  $H_{n,p}$  behaves and (4.3) holds. Fix any such choice and let  $t_i = |V(T_i)|$  for  $i = 1, \dots, N$ . To shorten coming formulae, let

$$\mathbb{P}_c\{\cdot\} = \mathbb{P}\{\cdot \mid C_{n,p}, T_1, \dots, T_N\},$$

and define  $\mathbb{E}_c$  and  $\text{Var}_c$  similarly. Given  $C_{n,p}$  and  $T_1, \dots, T_N$ , we may write  $X$  as

$$X = |B_1| + \sum_{i=1}^N t_i \mathbf{1}[T_i \text{ attaches to } B_1].$$

Since  $H_{n,p}$  behaves, by the estimates above,

$$\frac{|B_1|}{|V(C_{n,p})|} \geq \frac{(1 - \delta)\epsilon^2 n/9}{(2 + \delta)\epsilon^2 n} \geq \frac{1 - \delta}{18(1 + \delta)} \geq \frac{1}{22}.$$

Since the points of attachment of  $T_1, \dots, T_N$  in  $C_{n,p}$  are uniform and  $\sum_{i=1}^N t_i = |V(H_{n,p}) \setminus V(C_{n,p})|$ , it thus follows that

$$\mathbb{E}_c(X) = |B_1| + \frac{|B_1|}{|V(C_{n,p})|} \cdot |V(H_{n,p}) \setminus V(C_{n,p})| > \frac{\epsilon n}{22}, \quad (4.9)$$

the preceding inequality holding for  $n$  sufficiently large since  $H_{n,p}$  behaves. Next, given  $C_{n,p}$  and  $T_1, \dots, T_N$ ,  $|B_1|$  is determined and  $X - |B_1|$  is a sum of independent random variables  $t_i \mathbf{1}[T_i \text{ attaches to } B_1]$ ,  $i = 1, \dots, N$ . Hence,

$$\text{Var}_c(X) = \sum_{i=1}^N t_i^2 \text{Var}_c(\mathbf{1}[T_i \text{ attaches to } B_1]) \leq \sum_{i=1}^N t_i^2.$$

By Chebyshev's inequality, when  $n$  is large enough that (4.9) holds, we thus have

$$\mathbb{P}_c\left\{X < \frac{\epsilon n}{44}\right\} \leq \frac{\sum_{i=1}^N t_i^2}{(\epsilon n/44)^2}.$$

Since we have assumed that (4.3) holds, and that  $H_{n,p}$  behaves,

$$\sum_{i=1}^N t_i^2 \leq \max_{1 \leq i \leq N} t_i \cdot \sum_{i=1}^N t_i \leq C_{16} \varepsilon^{-2} \log(n\varepsilon^3) \cdot n\varepsilon$$

and thus, for  $n$  large enough,

$$\mathbb{P}_c \left\{ X < \frac{\varepsilon n}{44} \right\} \leq C_{17} \frac{n\varepsilon^{-1} \log(n\varepsilon^3)}{(\varepsilon n)^2} = C_{17} \frac{\log(n\varepsilon^3)}{n\varepsilon^3} \rightarrow 0$$

as  $n \rightarrow \infty$ . An identical argument yields the same lower bound with  $X$  equal to  $|B_2^+|$  or  $|B_3^+|$ . (We do not actually care about  $|B_2^+|$ .) This establishes  $(\star)$  and completes the proof.  $\square$

It was proved in [15] that when  $p = (1 + \varepsilon)/n$  with  $\varepsilon > 0$  and fixed, there is  $f(\varepsilon) > 0$  such that the diameter of  $H_{n,p}$  is  $f(\varepsilon) \log n + O_p(1)$ . Riordan and Wormald [15] have also proved tight bounds on the diameter of random graphs in the barely supercritical phase; these bounds can be used to yield upper bounds that complement (though do not quite match) the lower bound of Theorem 4.1. Riordan and Wormald prove that for  $\varepsilon = \varepsilon(n)$  with  $\varepsilon = o(1)$  and  $\varepsilon n^{1/3} \geq e^{(\log^* n)^4/3}$ , and  $p = (1 + \varepsilon)/n$ , w.h.p. the diameter of  $H_{n,(1+\varepsilon)/n}$  is  $\Theta(\log(\varepsilon^3 n)/\varepsilon)$ . (In fact, the result from [15] is substantially stronger and more precise than this.) From this bound, it follows immediately that for  $p$  in the above range, w.h.p.  $\sigma^{2*}(H_{n,p}) = O(\log^2(\varepsilon^3 n)/\varepsilon^2)$ . The upper bound is within a  $\log^2(\varepsilon^3 n)$  factor of the lower bound in Theorem 4.1. Note that when  $\varepsilon = e^{(\log^* n)^4/3}/n^{1/3}$ , this factor is just  $(\log^* n)^4$ , though as  $\varepsilon$  increases the upper bound becomes less tight.

## 5. LARGE DEGREES YIELD NEAR MINIMAL SPREAD

We saw that w.h.p. the random regular graph  $G(n, d)$  has bounded spread for any fixed  $d \geq 3$ , and similarly the random graph  $H_{n,c/n}$  has bounded spread for any fixed  $c > 1$ . It is easy to see that the complete graph  $K_n$  has spread  $1/4$  if  $n$  is even and  $1/4 - 1/(4n^2)$  if  $n$  is odd. This of course gives the minimum possible values of the spread.

The above suggests another natural question for random graphs. How large must degrees be for the spread to be close to  $1/4$ ? We shall see that for random regular graphs, what is needed is simply for the degree  $d$  to be big enough.

Firstly we note the deterministic result that the average degree must be large in order for the spread to be close to  $1/4$ , and then we give a matching result that for random regular graphs high degree is sufficient.

**Proposition 5.1.** *For any fixed  $d \geq 2$  there exists  $\delta > 0$  such that if the connected graph  $G$  has average degree at most  $d$  and  $|V(G)| \geq 3d$  then  $\sigma^{2*}(G) \geq 1/4 + \delta$ . (We can take  $\delta = 1/(6d)$ .)*

It is well known that for each fixed integer  $d \geq 3$ , w.h.p. the random graph  $G(n, d)$  is connected [5] and so we may talk of  $\sigma^{2*}(G(n, d))$ .

**Theorem 5.2.** *For each  $\varepsilon > 0$  there exists  $d_0$  such that for each  $d \geq d_0$  w.h.p.  $\sigma^{2*}(G(n, d)) < 1/4 + \varepsilon$ .*

*Proof of Proposition 5.1.* We shall show that if  $|G| = n$  then

$$\sigma^{2*}(G) \geq \frac{1}{4} + \left(\frac{1}{d} - \frac{2}{n}\right)\left(1 - \frac{1}{d}\right). \quad (5.1)$$

Note that this gives  $\sigma^{2*}(G) \geq 1/4 + \frac{1}{6d}$  if  $d \geq 2$  and  $n \geq 3d$ .

Let  $t = \lfloor \frac{n}{2d} \rfloor$ , let  $T$  consist of  $t$  vertices of least degree, and let  $U$  be the set of vertices adjacent to a vertex in  $T$ . Note that  $|U| \leq n/2$ . Let  $A \subseteq [n] \setminus T$  be such that  $A \supseteq U \setminus T$  and  $|A| = a := \lfloor \frac{n}{2} \rfloor$ . Let  $B = [n] \setminus (T \cup A)$ .

Let  $f(v) = 0$  on  $B$ , 1 on  $A$  and 2 on  $T$ . For  $X$  uniformly distributed over the vertices, and writing  $f$  for  $f(X)$ , we have  $\mathbb{E} f = (1/n)(a + 2t)$  and  $\mathbb{E} f^2 = (1/n)(a + 4t)$ , and hence

$$\begin{aligned} \text{Var}(f) &= (1/n)(a + 4t) - (1/n^2)(a^2 + 4at + 4t^2) \\ &= \frac{a}{n}\left(1 - \frac{a}{n}\right) + \frac{4t}{n} - \frac{2t}{n} + \frac{2t}{n^2}\mathbf{1}[n \text{ odd}] - \frac{4t^2}{n^2} \\ &= 1/4 - \frac{1}{4n^2}\mathbf{1}[n \text{ odd}] + \frac{2t}{n} + \frac{2t}{n^2}\mathbf{1}[n \text{ odd}] - \frac{4t^2}{n^2} \\ &\geq 1/4 + \frac{2t}{n}\left(1 - \frac{2t}{n}\right) \\ &\geq 1/4 + \left(\frac{1}{d} - \frac{2}{n}\right)\left(1 - \frac{1}{d}\right). \quad \square \end{aligned}$$

To prove Theorem 5.2 we need an expansion result for random regular graphs. Given  $\beta > 1$  and  $0 < \eta < 1$  let us say that a graph  $G = (V, E)$  has  $(\beta, \eta)$ -*expansion* if for each  $T \subset V$  with  $|T| \leq (1 - \eta)|V|/\beta$  we have  $|T \cup N(T)| \geq \beta|T|$ .

**Lemma 5.3.** *For each  $\beta > 1$  and  $0 < \eta < 1/2$  there exists  $d_0$  such that for all  $d \geq d_0$  w.h.p.  $G(n, d)$  has  $(\beta, \eta)$ -expansion.*

*Proof.* We consider the configuration model for  $G(n, d)$ . Let  $\alpha > 0$  ( $\alpha$  large). For a positive integer  $t$  let  $f_{n,d}(t)$  be the expected number of pairs  $T$  and  $U$  of sets of disjoint cells where  $|T| = t$  and  $|U| = u := \lfloor \alpha t \rfloor$ , and each neighbour of a stub in a cell in  $T$  is in  $T \cup U$ . Let  $t_0 = \lfloor (1 - \eta)|V|/\beta \rfloor$ . We aim to upper bound this quantity  $f_{n,d}(t)$ , in order to show that  $\sum_{t=1}^{t_0} f_{n,d}(t) = o(1)$ . The lemma will then follow, with  $\beta = 1 + \alpha$ .

Note first that, since  $\frac{d(t+u)-j}{dn-j} \leq \frac{t+u}{n}$  for each  $0 < j < dn$ , the probability that each neighbour of a stub in a cell in  $T$  is in  $T \cup U$  is at most  $(\frac{t+u}{n})^{dt/2}$ . (If we choose the neighbours of the  $dt$  stubs in cells in  $T$  first, we have to make at least  $dt/2$  such choices.) Hence

$$f_{n,d}(t) \leq \binom{n}{t} \binom{n}{u} \left(\frac{t+u}{n}\right)^{dt/2}$$

$$\begin{aligned}
&\leq \left(\frac{ne}{t}\right)^t \left(\frac{ne}{u}\right)^u \left(\frac{t+u}{n}\right)^{dt/2} \\
&\leq \left(\frac{ne}{t}\right)^t \left(\frac{ne}{\alpha t}\right)^{\alpha t} \left(\frac{(1+\alpha)t}{n}\right)^{dt/2} \\
&= \left(e^{1+\alpha} \alpha^{-\alpha} (1+\alpha)^{d/2} t^{d/2-1-\alpha} n^{1+\alpha-d/2}\right)^t \\
&= \left(e^{1+\alpha} \alpha^{-\alpha} (1+\alpha)^{1+\alpha} \left(\frac{(1+\alpha)t}{n}\right)^{d/2-1-\alpha}\right)^t.
\end{aligned}$$

Now  $\alpha^{-\alpha}(1+\alpha)^{1+\alpha} = (1+\alpha)(1+1/\alpha)^\alpha \leq (1+\alpha)e$ . So

$$f_{n,d}(t) \leq \left((1+\alpha)e^{2+\alpha} \left(\frac{(1+\alpha)t}{n}\right)^{d/2-1-\alpha}\right)^t.$$

Let  $\alpha > 0$  be sufficiently large that  $\ln(1+\alpha)+2+\alpha \leq 2\alpha$ . Let  $d_0 \geq 6(1+\alpha)$ , so that  $d/2 - 1 - \alpha \geq d/3$  when  $d \geq d_0$ . For such  $d$

$$f_{n,d}(t) \leq \left(e^{2\alpha} \left(\frac{(1+\alpha)t}{n}\right)^{d/3}\right)^t.$$

If  $1 \leq t \leq \ln^2 n$  say then

$$f_{n,d}(t) \leq \left(e^{2\alpha} \left(\frac{(1+\alpha)\ln^2 n}{n}\right)^{d/3}\right)^t = O(1/n),$$

since  $d \geq 6$ . Also, since  $\frac{(1+\alpha)t}{n} \leq 1 - \eta \leq e^{-\eta}$ , for  $1 \leq t \leq t_0$  we have

$$f_{n,d}(t) \leq \left(e^{2\alpha} e^{-\eta d/3}\right)^t.$$

From these bounds it is easy to complete the proof, with  $\beta = 1 + \alpha$ .  $\square$

**Lemma 5.4.** *Let  $\beta \geq 3$ ,  $\eta = \beta^{-1}$  and  $n \geq 6\beta + \beta^2/2$ , and let  $G = (V, E)$  have  $(\beta, \eta)$ -expansion. Let  $f$  be an integer-valued function on  $V$  with median 0. Let  $V_{\geq i}$  denote  $\{v \in V : f(v) \geq i\}$  and so on. Assume that  $|V_{\geq 1}| \geq |V_{\leq -1}|$ . Then*

$$|V_{\geq i}| \leq \beta^{-(i-1)} n/2 \text{ and } |V_{\leq -i}| \leq 2\beta^{-i} n \text{ for each } i \geq 1. \quad (5.2)$$

*Proof.* Note that  $|V_{\leq 0}| \geq n/2$  and  $|V_{\geq 0}| \geq n/2$ . Observe also that  $N(V_{\geq i}) \subseteq V_{\geq i-1}$ . If  $|V_{\geq 2}| > (1-\eta)n/\beta$  then, choosing a set  $T \subset V_{\geq 2}$  with  $|T| = \lfloor (1-\eta)n/\beta \rfloor$ ,

$$|V_{\geq 1}| \geq |T \cup N(T)| \geq \beta|T| > (1-\eta)n - \beta \geq n/2 \geq |V_{\geq 1}|, \quad (5.3)$$

a contradiction: thus  $|V_{\geq 2}| \leq (1-\eta)n/\beta$ . Hence  $|V_{\geq 2}| \leq \frac{1}{\beta}|V_{\geq 1}| \leq \frac{n}{2\beta}$ , and further for all  $i \geq 1$  we have  $|V_{\geq i}| \leq \beta^{-(i-1)}|V_{\geq 1}|$ . Similarly, for all  $i \geq 1$  we

have  $|V_{\leq -i}| \leq \beta^{-(i-1)}|V_{\leq -1}|$ . Hence it suffices to show (5.2) for  $i = 1$ , i.e., that  $|V_{\geq 1}| \leq n/2$ , which is trivial, and  $|V_{\leq -1}| \leq 2n/\beta$ .

Recall that  $|V_{\geq 1}| \geq |V_{\leq -1}|$ . We consider two cases, depending on the size of  $V_{\geq 1}$ . If  $|V_{\geq 1}| \leq (1 - \eta)n/\beta$  then  $|V_{\geq -1}| \leq |V_{\geq 1}| < 2n/\beta$ . If  $|V_{\geq 1}| > (1 - \eta)n/\beta$  then  $|V_{\geq 0}| \geq (1 - \eta)n - \beta$  as in (5.3), so  $|V_{\leq -1}| \leq \eta n + \beta \leq 2n/\beta$ .  $\square$

The last lemma easily yields:

**Lemma 5.5.** *For any  $\epsilon > 0$  there exists  $\beta > 1$  such that each graph  $G$  with  $(\beta, \beta^{-1})$ -expansion and  $n$  large enough satisfies  $\sigma^{2*}(G) < 1/4 + \epsilon$ .*

*Proof.* Let  $f$  be an integer-valued Lipschitz function on  $G$ . We may assume that the median of  $f$  is 0, and (by symmetry) that  $|V_{\geq 1}| \geq |V_{\leq -1}|$ . Then Lemma 5.4 yields, if  $\beta \geq 3$  and  $n$  is large,

$$\text{Var}(f) \leq \mathbb{E} \left| f - \frac{1}{2} \right|^2 \leq \frac{1}{4} + \sum_{i \neq 0,1} \frac{|V_i|}{n} (i - 1/2)^2 \leq \frac{1}{4} + O(\beta^{-1}). \quad \square$$

Lemmas 5.3 and 5.5 complete the proof of Theorem 5.2.

## 6. OPEN PROBLEMS

We saw in the preceding section that high degree is precisely what is needed to force the spread of  $G(n, d)$  close to  $1/4$ . We believe that a corresponding result should hold for  $G_{n,c/n}$ .

**Problem 6.1.** Is it the case that for each  $\epsilon > 0$ , there exists  $c_0$  such that for each  $c \geq c_0$  w.h.p.  $\sigma^{2*}(H_{n,c/n}) < 1/4 + \epsilon$ ?

If Theorem 3.1 holds uniformly (in the sense that for any  $c > 1$ ,  $\alpha = \alpha(c)$  can be chosen such that the conclusions of the theorem hold in  $H_{n,c'/n}$  for all  $c' \geq c$ , with this value of  $\alpha$ ) then the proof of Theorem 3.3 can be modified to yield an affirmative answer to the above question. This uniformity seems very likely to hold, but does not follow immediately from the proof of Theorem 3.1 given in [2].

Corollary 2.2 suggests that the spread of  $G(n, d)$  might converge (in probability) to a constant, and similarly for Corollary 3.4 and  $H_{n,c/n}$ .

**Problem 6.2.** Do there exist constants  $\alpha_d$  for each  $d \geq 3$  and  $\beta_c$  for each  $c > 1$  such that  $\sigma^{2*}(G(n, d)) \xrightarrow{P} \alpha_d$  and  $\sigma^{2*}(H_{n,c/n}) \xrightarrow{P} \beta_c$  as  $n \rightarrow \infty$ ?

We know that if the constants  $\alpha_d$  exist then they are (weakly) decreasing in  $d$  and tend to  $1/4$  as  $d \rightarrow \infty$ . It seems likely that the analogous results should hold for  $G_{n,c/n}$ .

**Problem 6.3.** If the constants  $\beta_c$  exist, are they decreasing in  $c$ , and do they tend to  $1/4$  as  $c \rightarrow \infty$ ?



## REFERENCES

- [1] N. Alon, R. Boppana, and J. Spencer. An asymptotic isoperimetric inequality. *Geometric and Functional Analysis*, 8(3):411–436, 1998.
- [2] I. Benjamini, G. Kozma, and N. Wormald. The mixing time of the giant component of a random graph. arXiv:0610459v1 [math.PR], 2006.
- [3] S. Bobkov, C. Houdré, and P. Tetali.  $\lambda_\infty$ , vertex isoperimetry and concentration. *Combinatorica*, 20(2):153–172, 2000.
- [4] B. Bollobás. *Random Graphs*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, second edition, 2001.
- [5] B. Bollobás. Random graphs. In *Combinatorics, Proceedings, Swansea 1981*, volume 52 of *London Math. Soc. Lecture Note Series*, pages 80–102. Cambridge University Press, 1981.
- [6] S. Janson and M.J. Luczak. A simple solution to the k-core problem. *Random Structures and Algorithms*, 30(1-2):50, 2007.
- [7] S. Janson, D.E. Knuth, T. Luczak, and B. Pittel. The birth of the giant component. *Random Structures and Algorithms*, 4(3):233–359, 1993.
- [8] S. Janson, T. Luczak, and A. Ruciński. *Random Graphs*. Wiley, New York, 2000.
- [9] V. F. Kolchin. *Random Mappings*. Optimization Software Inc., 1986.
- [10] T. Luczak. Size and connectivity of the k-core of a random graph. *Discrete Mathematics*, 91(1):61–68, 1991.
- [11] T. Luczak. Component behavior near the critical point of the random graph process. *Random Structures and Algorithms*, 1(3):287–310, 1990.
- [12] T. Luczak. Cycles in a random graph near the critical point. *Random Structures and Algorithms*, 2:421–440, 1991.
- [13] Y.L. Pavlov. The asymptotic distribution of maximum tree size in a random forest. *Theory of Probability and its Applications*, 22:509, 1978.
- [14] Yu. L. Pavlov. *Random Forests*. VSP, 2000.
- [15] O. Riordan and N. Wormald. The diameter of sparse random graphs. arXiv:0808.4067v1 [math.PR], September 2008.
- [16] N. Wormald. The asymptotic connectivity of labelled regular graphs. *Journal of Combinatorial Theory, Series B*, 31(2):156–167, 1981.

DÉPARTEMENT DE MATHÉMATIQUES ET DE STATISTIQUE, UNIVERSITÉ DE MONTRÉAL,  
CP 6128, SUCC. CENTRE-VILLE, MONTRÉAL, QUÉBEC, H3C 3J7, CANADA

*E-mail address:* `louigi@gmail.com`

*URL:* `http://www.dms.umontreal.ca/~addario/`

DEPARTMENT OF MATHEMATICS, UPPSALA UNIVERSITY, PO Box 480, SE-751 06  
UPPSALA, SWEDEN

*E-mail address:* `svante.janson@math.uu.se`

*URL:* `http://www.math.uu.se/~svante/`

MATHEMATICAL INSTITUTE, UNIVERSITY OF OXFORD, 24–29 ST GILES', OXFORD  
OX1 3LB, UK

*E-mail address:* `cmcd@stats.ox.ac.uk`

*URL:* `http://www.stats.ox.ac.uk/~cmcd/`